Stereo Matching in the Presence of Narrow Occluding Objects Using Dynamic Disparity Search

Umesh R. Dhond and J. K. Aggarwal

***Abstract*—Most contemporary stereo correspondence algorithms impose global consistency among candidate match-points using Spatial Hierarchy Mechanism- (SHM) based techniques that rely on either the local support within a 2D neighborhood in the image plane and/or cooperative processes between multiple levels of a pixel-resolution or structural-description hierarchy. We analyze the stereo matching failures in SHM-based techniques in the presence of narrow occluding objects and propose the Dynamic Disparity Search (DDS) framework to reduce false-positive matches. Experiments with indoor and outdoor scenes demonstrate a significant reduction in the false-positive match rates of a DDS-based stereo algorithm as compared to those of two existing algorithms.**

***Index Terms*—Stereo, matching, correspondence, image analysis, binocular, occlusion, shadow region, disparity pool, 3D structure, triangulation, dynamic disparity search.**

I. Introduction

The central problem in the stereo analysis paradigm is to find the correspondence between two or more images captured from multiple viewpoints. The correspondence is achieved in two matching stages—local and global. The local matching stage finds the possible candidate matches within a given interval of allowable disparity values based upon the epipolar search constraint and the geometrical similarity constraint. The global matching stage is responsible for imposing global consistency among the candidate match points by disambiguating the multiple candidate matches and avoiding false-positives. Most conventional global matching techniques [1] use

1) the local support provided by pixels within a 2D neighborhood in the image plane and/or
2) cooperative processes between multiple levels of a pixel-resolution hierarchy or a 2D structural-description hierarchy

and are collectively referred to as Spatial Hierarchy Mechanism (SHM)-based techniques.

This correspondence analyzes stereo mismatches in scenes containing narrow occluding objects when using SHM-based techniques alone in the global matching stage. Based upon the lessons learnt from this analysis, we propose a new global matching framework—*Dynamic Disparity Search* (DDS)—which reduces the number of stereo mismatches in scenes containing narrow occluding objects. The DDS-based framework extends the conventional SHM-based framework by

1) splitting the entire allowable disparity range into two (or more) disjoint disparity pools,
2) dynamically varying the allowable disparity range in each pool from the narrowest to the widest intervals, and
3) propagating stereo disparities through the spatial hierarchy as well as the disparity hierarchy mechanisms.

As an experiment, we extend the Marr-Poggio-Grimson (MPG) algorithm [2] using the DDS framework and compare the False-Positive Rates (FPRs) of the *DDS-based implementation of MPG algorithm* (henceforth referred to as the DDS algorithm) to that of the MPG algorithm as well as to the multiple-baseline Okutomi-Kanade-Nakahara (OKN) [3] algorithm.

The rest of this correspondence is organized as follows: Section II briefly discusses the global matching constraints used by SHM-based techniques. The stereo mismatches caused by relying on SHM-based global matching constraints alone are discussed in Section III. Section IV describes the design and implementation of the DDS-based extension of the MPG algorithm. Experimental results are presented in Section V and concluding remarks in Section VI.

Manuscript received Aug. 19, 1993; revised Feb. 6, 1995.

U.R. Dhond is with Schlumberger Austin Systems Center, 8311 N. FM 620 Road, Austin, TX 78726. email: udhond@slb.com.

J.K. Aggarwal is with Computer and Vision Research Center, Department of Electrical and Computer Engineering, ENS 520, The University of Texas at Austin, Austin, TX 78712-1084, USA. email: jka@emx.cc.utexas.edu.

IEEECS Log Number P95070.

II. Previous Work: Global Matching Constraints

A. Spatial Hierarchy Mechanism

Conceptually, the Spatial Hierarchy Mechanism (SHM) for global matching is based on the assumption that the 3D world is composed of cohesive matter. It imposes global consistency among stereo disparities using

1) a 2D local support neighborhood in the image plane and/or
2) cooperative processes between a multilevel pixel-resolution hierarchy or a structural-description hierarchy.

The phrase "using a 2D local support neighborhood," refers to the entire spectrum ranging from the simple neighborhood disparity-averaging technique to the higher-order models that capture disparity variation within a 2D neighborhood window [4], [5], [6]. The presence of shadow regions or occlusion boundaries within this 2D neighborhood introduces errors in the local support computation as detailed in Section III. The following are eight of the many SHM-based techniques described in the literature [1]: disparity smoothness along figural contours [7], the disparity gradient limit [8], the minimal differential disparity constraint for line segments [9], hierarchical structural descriptions (surfaces, junctions, edges) [10], [11], the preservation of L-to-R feature ordering [12], [10], [13], relaxation labeling [14], [15], [16], surface fitting techniques [5], and hierarchical graph-matching techniques [17].

III. Occlusion Analysis

This section discusses various scenarios of stereo mismatches caused by narrow occluding objects when a particular SHM-based technique, disparity-averaging SHM, is used for global matching. First, we define narrow occluding objects based on two sufficient conditions [18].

A. Definition

Let w be the pixel width of an occluding object A in the left image of a stereo image pair. Let w_N denote the width of the local support neighborhood used by a stereo algorithm for disambiguation of multiple candidate disparities. Also, let d_{BG} and d_{FG} be the average stereo disparities of the occluded (background) and occluding (foreground) regions in the left image, respectively. If either

1) $w < w_N$ and/or
2) $w < (d_{FG} - d_{BG})$, then A is a narrow occluding object.

In other words, an occluding object is termed as narrow if either

1) all the object's image pixels violate the disparity smoothness constraint or
2) its boundary pixels with the occluded object violate the L-to-R ordering constraint.

B. Shadow Regions

Shadow regions existing in the left image have no true matches in the right image, and vice versa. Consider the scene with narrow occlusions in Fig. 1 where a smooth object in the background is occluded by narrow occluding objects in the foreground. For any point X, $avg(N(X))$ computes the average disparity value of all uniquely matched pixels within a given neighborhood $N(X)$. Contributions to $avg(N(X))$ consist of the percentages of correctly matched background (BG) and foreground (FG) features within $N(X)$, β and ϕ, and those of erroneously matched pixels ϵ. Further, if $d_\beta(N(X))$, $d_\phi(N(X))$, and $d_\epsilon(N(X))$ denote the true average disparity value for each of the three sets above, the $avg(N(X))$ will equate to,

$$avg(N(X)) = (d_\beta(N(X)) \times \beta) + (d_\phi(N(X)) \times \phi) + (d_\epsilon(N(X)) \times \epsilon). \quad (1)$$

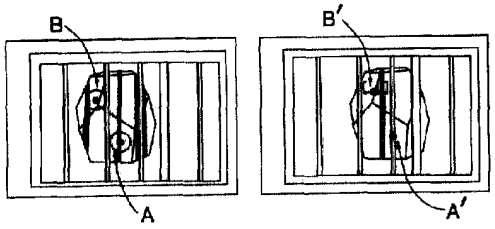


Fig. 1. Mismatches due to shadow regions: Point A lies within a shadow region, whereas B's local support neighborhood intersects a shadow region.

For the shadow-region feature A in Fig. 1, all $N(A)$ disparities vary smoothly since they all belong to the same 3D surface. To isolate the effects of shadow regions, let's assume no mismatched pixels exist within $N(A)$) at the outset. Then, $\phi = 0$ and $\beta >> \epsilon \geq 0$. Consequently, the regional disparity average computed by (1) will be close to d_β.

If the SHM-based algorithm uses the disparity estimate $avg(N(A))$ to disambiguate a set of multiple candidate matches, the unique match so obtained is *always* a false-positive since the point A belongs to a shadow region. In other cases, when the SHM transmits the disparity estimate $avgN(A))$ from a coarser to a finer resolution, it propagates an incorrect disparity estimate to the finer resolution. Plus, unlike the occlusions analyzed in Chung-Nevatia [11] and Little-Gillett [19], the positions of shadow regions cannot be localized for narrow occlusions since they violate the L-R ordering constraint.

C. Shadow Drift

In Fig. 1, the mismatched shadow-region pixels in the circular support neighborhood $N(B)$ contribute noisy disparity values to the d_ϵ factor in (1) for the BG feature point B. Assuming no other mismatched pixels exist in $N(B)$, ϵ represents the percentage of matched shadow-region pixels, which is scene-dependent and could be as high as 0.5. This noise causes the disparity estimate of $N(B)$ to drift from its true value and the phenomenon is termed as *Shadow Drift*.

D. Occlusion Boundaries

Fig. 2 shows a point C on a narrow occluding object. Since C is a narrow occluding feature point, $N(C)$ is weighed in favor of the average background disparity value d_β due to insufficient support from neighboring foreground (occluding) matched feature points. Thus, even though the true match of C is C′, as shown in Fig. 2, $avg(N(C))$ tends towards the disparity associated with C″ because $\phi << \beta$ since C belongs to a narrow occlusion. Such incorrect weighing of local support is the key reason for foreground narrow occluding features being matched to background feature points in the neighborhood of a narrow occlusion boundary.

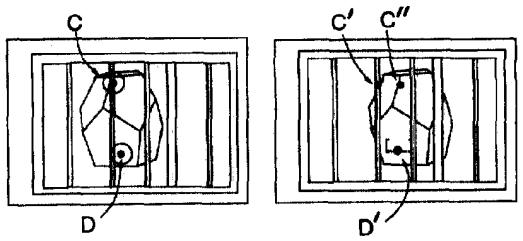


Fig. 2. Mismatches due to occlusion boundaries: Point C lies on an occlusion boundary, whereas D's local support neighborhood intersects an occlusion boundary.

E. Occlusion Drift

Consider a background feature point D in Fig. 2 whose local support neighborhood contains a narrow occlusion boundary. Depending upon the relative proportion of the areas occupied by the foreground and background regions within the local support neighborhood $N(D)$, the neighborhood disparity estimate $avg(N(D))$ returned by (1) drifts closer to or farther away from the true match D′.

As the above four scenarios indicate, the relative percentages of the occluded and occluding object pixels affect the local disparity average returned by the SHM-based disparity smoothness constraint and cause matching failures. Conversely, if matching is performed in two separate disjoint disparity pools—background and foreground—the two would cease to interfere in the local support neighborhoods of pixels belonging to the other disparity pools. Further, if the range of allowable disparities is narrow during the initial matching stages, fewer number of candidate matches will enter the disambiguation stage, and hence fewer shadow region pixels will get matched. The two basic principles of the DDS framework—the use of multiple disjoint disparity pools and dynamic variation of the range of allowable disparities—were motivated by the analysis [18] of the four stereo mismatch scenarios presented in this section. The next section contains the detailed design and implementation of the DDS-based stereo matching algorithm extended from the Marr-Poggio-Grimson SHM-based technique. Any existing SHM-based stereo algorithm can be extended using the DDS framework.

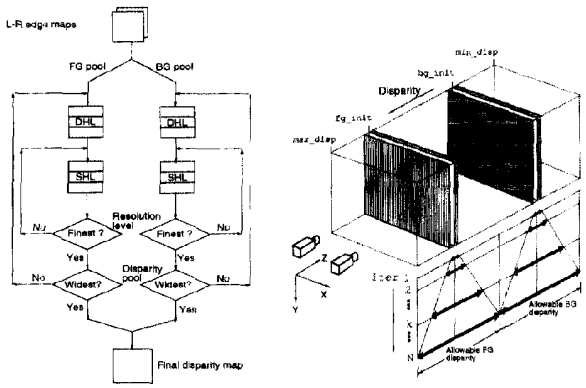
IV. Dynamic Disparity Search-Based Algorithm

The DDS framework is a general design methodology which defines a family of stereo algorithms based on various design choices of imaging configurations, matchable features, local matching constraints, and a spatial hierarchy mechanism for the global matching stage. Without loss of generality, the spatial hierarchy mechanism chosen for our DDS-based algorithm design and implementation is based on the MPG algorithm [2].

Fig. 3(a) draws the control flow-chart of the overall DDS global matching technique. We use two disparity pools: background (BG) and foreground (FG). The control flow for each of the BG and FG disparity pools consists of two nested loops, the *Disparity Hierarchy Loop* (DHL) and the *Spatial Hierarchy Loop* (SHL), which impose the disparity and spatial hierarchy constraints, respectively.

A. DHL Initialization

The entire allowable disparity range [min_disp, max_disp] is partitioned into two equal disjoint ranges, background (BG) and foreground (FG), as Fig. 3(b) shows. Initialize the disparity maps of the BG and FG pools with the midpoints of the BG and FG disparity ranges, bg_init and fg_init, respectively. Set the allowable disparity range of the BG and FG pools during the DHL's first iteration, iter 1 in Fig. 3(b), to [bg_init-1, bg_init+1] and [fg_init-1, fg_init+1], respectively. In effect, during iter 1, the DHL allows matches to occur only for pixels whose disparity lies within a very narrow interval. During each successive DHL iteration, the individual, allowable FG and BG disparity pools are widened until they span the entire allowable disparity range [min_disp, max_disp], as Fig. 3(b) shows. The pixels with disparities belonging to the FG and BG pools are manipulated separately in each of the disparity hierarchy loops. At any given time, the disparity maps of pixels within each of the two disparity pools FG and BG are maintained separately and denoted by fg_dmap and bg_dmap, respectively.



(a) (b)

Fig. 3. Dynamic Disparity Search mechanism: (a) control flow diagram, (b) dynamic disparity pools FG and BG.

B. SHL Initialization

The SHL is initialized during each DHL iteration. The SHL computations are dependent on the particular SHM technique that is being generalized within the DDS framework. The SHM in the MPG algorithm used in our implementation requires a coarse-to-fine spatial resolution hierarchy. The L and R edge maps of the coarsest resolution are loaded into the local matching module for each disparity pool, BG and FG. Each feature point is assigned an initial disparity estimate, d_i. During the SHL initialization of iter 1 of the DHL, the initial disparity estimates for the BG and FG pools are bg_init and fg_init, respectively. For the SHL initialization in kth DHL ($1 < k \leq N$), the disparity map obtained at the finest spatial resolution of the SHL belonging to the (k - 1)th (previous) DHL is used.

C. Stereo Matching Computations

Stereo matching is performed separately for each BG and FG pool at each spatial resolution in three stages:

1) For each feature point P_L in the left image, find a list of candidate matches. Retain only those candidates that lie within the allowable disparity ranges for the BG and FG pools.
2) Identify the uniquely matched feature points within each disparity pool BG and FG.
3) Disambiguate multiple matches using the SHM-specific global matching constraint. The only difference in the DDS framework is that the global matching constraint is applied *separately* in each of the disjoint disparity pools. Pixels with no unique match are labeled *NOMATCH*.

D. Spatial Hierarchy Loop Termination

The SHL termination computations are entirely dependent on the choice of spatial hierarchy mechanism being implemented within the DDS framework. This section describes the SHL termination conditions used for the DDS-based extension of the MPG spatial hierarchy mechanism. The SHL termination computation propagates the disparity map available at the end of the previous SHL iteration separately for each disparity pool, BG and FG.

If the current spatial resolution is not the finest, initialize the disparity map of the next finer resolution using a local neighborhood averaging of the unique disparities computed by the current SHL iteration as per the MPG-based SHM [2]. Alternately, if the current spatial resolution is the finest, prepare for the next DHL iteration, that is, check the DHL termination condition and perform DHL and SHL initializations.

E. Disparity Hierarchy Loop Termination

Check whether the disjoint disparity pools have reached the widest possible size. If the union of the disparity pools spans the entire allowable disparity range [min_disp, max_disp], the DHL is terminated. Otherwise, the width of the allowable disparity ranges of each BG and FG disparity pool is increased by a predetermined value.

F. Composite Disparity Map

The DHL computation returns a disparity map corresponding to each disparity pool—BG and FG—at the finest spatial resolution. The disparity maps corresponding to the two disparity pools are merged with a union operation. If $d_b(i,j)$ and $d_f(i,j)$ are the BG and FG disparity maps, the composite disparity map $d_{comp}(i,j)$ is,

$$d_{comp}(i,j)=\begin{cases} d_b(i,j) & \text{if } d_f(i,j)=NOMATCH \\ d_f(i,j) & \text{otherwise.} \end{cases}$$

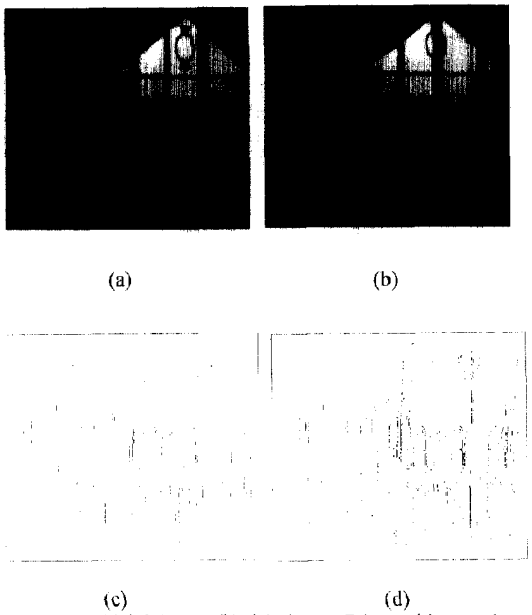
V. Experimental Results

This section quantifies the improvements in the stereo matching performance of a conventional SHM-based stereo matching algorithm redesigned using the proposed DDS global matching framework. The DDS-based stereo algorithm implementation described in Section IV is based on the local matching technique and spatial hierarchy mechanism of the Marr-Poggio-Grimson (MPG) algorithm [2]. This section describes stereo matching experiments comparing the performance of the DDS algorithm described in Section IV with the MPG algorithm itself. Next, we compare the above DDS implementation with the multiple baseline, trinocular stereo algorithm recently proposed by Okutomi, Kanade, and Nakahara (OKN) [3]. We evaluate the performance of each stereo algorithm by comparing the stereo disparity estimate obtained as a result of stereo matching to the ground truth.

Establishing the ground truth is an extremely important part of this evaluation process. No standard stereo datasets with precomputed ground truths are available. Our experimental results with real scenes use ground truth computed from the raw images. The entire image is segmented manually into convex polygons. The disparity map within the interior of each polygon is assumed to be planar. The true disparity is obtained manually at each polygon vertex from the stereo intensity images. The disparity along the edges and within the interior

of each polygon is obtained by linear interpolation. The False-Positive Rates (FPRs) reported in this section are computed as the ratio of the number of mismatched pixels to the total number of matched (correctly plus incorectly) pixels.

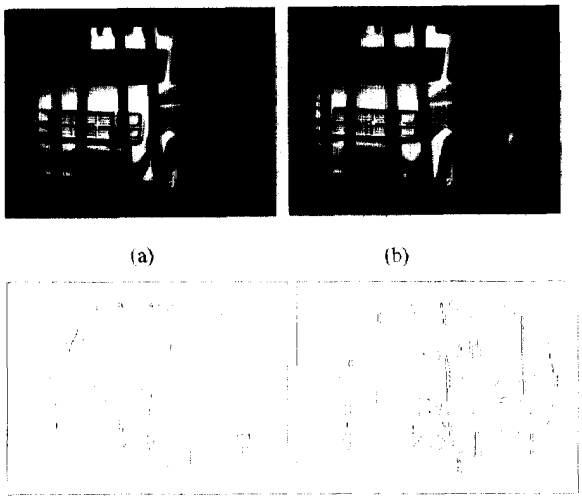
First, we compare the performance of the MPG and the DDS algorithms on two indoor scenes—HouseII and Truck. This is followed by comparisons of three outdoor scenes—Parking meters, Anna Hiss, and Parking lot—where we compare the MPG, DDS, and OKN algorithm results. The complete set of experimental results, which consists of two synthetic stereo data sets, five indoor, and five outdoor scenes, can be found in [18].



(a) (b)

(c) (d)

Fig. 4. House II: (a) left image, (b) right image. False-positive matches with (c) DDS, and (d) MPG techniques.



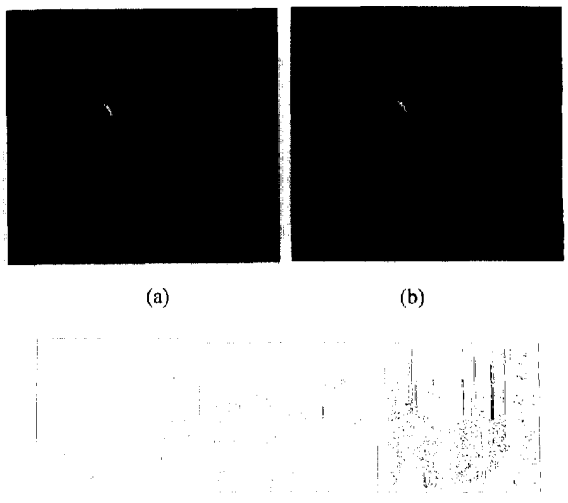
(a) (b)

(c) (d)

Fig. 5. Truck: (a) left image, (b) right image. False-positive matches with (c) DDS, and (d) MPG techniques.

The "House II" data set in Fig. 4 is an indoor laboratory scene of a toy house behind a wooden fence with a disparity range of (9, 38) pixels. The MPG algorithm results in several mismatches amongst background (window) and foreground (occluding bar) pixels. The "Truck" real indoor scene also has a nonuniform background disparity, as Fig. 5(a)–(b) illustrates. The images have a disparity range of (7, 20) pixels and show examples of the violation of the ordering constraint and the presence of shadow regions. The MPG results in Fig. 5(d) clearly show several incorrectly matched background shadow region pixels at the truck's rear end.

The following three sets of stereo data show outdoor scenes that were captured with multiple baselines to facilitate comparison with the Okutomi-Kanade-Nakahara (OKN) algorithm. The "Parking meters" dataset was obtained from Professor Takeo Kanade, at Carnegie-Mellon University. The latter two—Anna Hiss and Parking lot—were captured within the University of Texas-Austin campus. The "Parking meters" scene (Fig. 6), with a receding wall in the back and bushes and parking meters in front, possesses a disparity range of (2, 15) pixels. Most of the mismatches occur among the bushes which have a high feature density. The DDS matching algorithm still results in a modest FPR of 5.3% (MPG 13.4%; OKN 34.2%). The "Anna-Hiss" outdoor scene (Fig. 7) shows the entrance to the Anna Hiss gymnasium on the University of Texas campus with curved railings in the foreground occluding the entrance gate. It contains examples of order reversal and shadow regions which are typical characteristics of narrow occlusions. The "Parking lot" outdoor scene (Fig. 8) shows four cars partially occluded by two parking signs and a flood-light pole in the Department of Electrical and Computer Engineering, U.T. Austin, parking lot. The DDS matching results have an FPR of 4.2% (MPG 10.7%; OKN 28.0%). Table I enumerates all the stereo data sets with their image sizes, disparity ranges, and the allowable error tolerance used in computing the FPR. Table II summarizes the comparative matching performance of the DDS, MPG, and OKN algorithms.



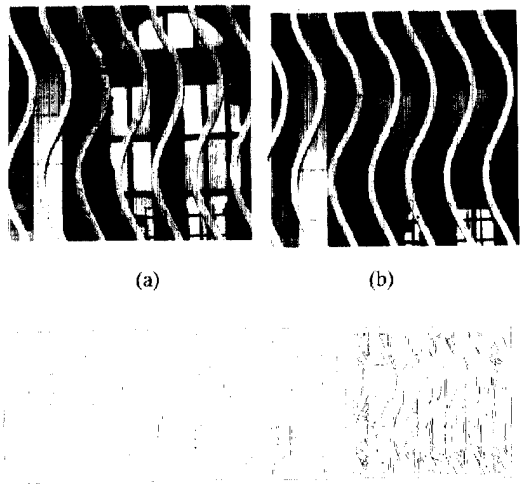
(a) (b)

(c) (d) (e)

Fig. 6. Parking meters: (a) left image, (b) right image. False-positive matches with (c) DDS, (d) MPG, and (e) OKN techniques. (Raw images courtesy Prof. Takeo Kanade, CMU.)

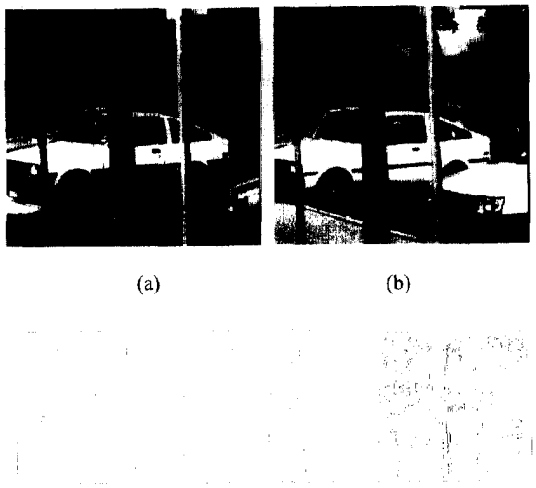
VI. Conclusion

We present a novel computational framework for stereo analysis that addresses the narrow occlusion problem. We analyze the scenarios of stereo matching failure of the conventional Spatial Hierarchy Mechanism (SHM) in the presence of narrow occluding objects. The occlusion analysis has motivated the proposed Dynamic Disparity Search (DDS) framework based on three fundamental guidelines:



(a) (b) (c) (d) (e)

Fig. 7. Anna-Hiss gym. (a) left image, (b) right image. False-positive matches with (c) DDS, (d) MPG, and (e) OKN techniques.



(a) (b) (c) (d) (e)

Fig. 8. Parking lot: (a) left image, (b) right image. False-positive matches with (c) DDS, (d) MPG, and (e) OKN techniques.

1) perform stereo matching using multiple disjoint disparity pools spanning the entire allowable disparity interval;
2) dynamically vary the allowable disparity interval within each disparity pool from the narrowest to the widest; and
3) propagate the stereo disparities using any conventional spatial hierarchy mechanism within each level of the disparity hierarchy.

Any conventional spatial hierarchy mechanism-based stereo algorithm can be redesigned using the above DDS-framework. As an illustrative example, we convert the conventional MPG [2] algorithm into its DDS-based version and find substantial reductions in the FPR.

TABLE I
SUMMARY OF STEREO DATA SETS

Scene	Size (Cols × rows)	Disparity range	Error Tolerance
House II	350 × 317	9 – 38	3
Truck	370 × 285	7 – 20	3
Prkg.Mtr.	257 × 241	2 – 15	3
Anna Hiss	256 × 240	2 – 15	3
Prkg. Lot	256 × 240	2 – 19	3

TABLE II
COMPARISON OF FALSE-POSITIVE RATES
(TOTAL MISMATCHED/TOTAL MATCHED AS A PERCENTAGE)

Scene	SHM	OKN	DDS
House II	1641/3945 = 41.6	–	696/4579 = 15.2
Truck	1087/5147 = 21.1	–	395/5051 = 7.8
Prkg.Mtr.	720/5387 = 13.4	3478/10162 = 34.2	240/4507 = 5.3
Anna Hiss	476/3043 = 15.6	3644/7586 = 48.0	166/4355 = 3.8
Prkg. Lot	241/2244 = 10.7	2087/7444 = 28.0	119/2815 = 4.2

Further research needs to be done to evaluate the impact of the DDS stereo matching framework on matching accuracy. One approach would be to study effects of redesigning other spatial-hierarchy mechanism based algorithms, like description-hierarchy ([10], [11]) or surface-fitting ([5]), using the DDS framework. Another method of doing comparative performance testing would be to compare the basic DDS-based MPG algorithm described in this paper with other existing algorithms like the comparison with the OKN [3] algorithm presented here.

Acknowledgments

This research was supported in part by Schlumberger and the Army Research Office under contract number DAAL-03-91-G-0050 and DAAH-04-94-G-0417. Also, many thanks to Dr. Lisa Judge for several suggestions that enhanced the readability of this manuscript.

References

[1] U.R. Dhond and J.K. Aggarwal, "Structure from stereo: A review," *IEEE Trans. Sys. Man Cyber.* vol. 19, pp. 1,489–1,510, Nov./Dec. 1989. Also in S.S. Iyengar and A. Elfes, eds., *Autonomous MobileRobots: Perception, Mapping, and Navigation*, vol. 1. Los Alamitos, Calif.: IEEE CS Press, 1991, pp. 25–46.

[2] W.E.L. Grimson, "A computer implementation of a theory of human stereo vision," *Phil. Trans. Royal Soc. London*, vol. B292, 1981, pp. 217–253.

[3] T. Tanade, M. Okutomi, and T. Nakahara, "A multiple-baseline stereo method," *Proc. DARPA Img. Understg. Wkshp.*, Jan. 1992.

[4] T.E. Boult and L.H. Chen, "Analysis of two new stereo algorithms," *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, 1988, pp. 177–182.

[5] W. Hoff and N. Ahuja, "Surfaces from stereo: Integrating feature matching, disparity information, and contour detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989.
[6] R.D. Eastman and A.M. Waxman, "Using disparity functionals for stereo correspondence and surface reconstruction," *Comp. Vision, Graphics, and Image Proc.*, vol. 39, July 1987, pp. 73–101.
[7] J.E.W. Mayhew and J.P. Frisby, "Psychophysical and computational studies towards a theory of human stereopsis," *Artificial Intelligence*, vol. 17, 1981, pp. 349–385.
[8] S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby, "Pmf: A stereo correspondence algorithm using a disparity gradient limit," *Perception*, vol. 14, pp. 449–470, 1981.
[9] G.G. Medioni and R. Nevatia, "Segment-based stereo matching," *Comp. Graphics Img. Proc.*, vol. 13, pp. 257–269, 1980.
[10] H.S. Lim and T.O. Binford, "Stereo correspondence: A hierarchical approach," *Proc. DARPA Img. Understg. Wkshp.*, Los Angeles, pp. 234–241, Feb. 1987.
[11] R.C.K. Chung and R. Nevatia, "Use of monocular groupings and occlusion analysis in an hierarchical stereo system," *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, Maui, Hawaii, pp. 50–55, June 1991.
[12] Y. Ohta and T. Kanade, "Stereo by intra- and inter-scanline search," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, pp. 139–154, Mar. 1985.
[13] H.H. Baker and T.O. Binford, "Depth from edge and intensity based stereo," *Proc. Seventh Int'l Joint Conf. AI*, pp. 631–636, Aug. 1981.
[14] S.T. Barnard and W.B. Thompson, "Disparity analysis of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, pp. 333–340, July 1980.
[15] Y.C. Kim and J.K. Aggarwal, "Positioning 3-d objects using stereo images," *IEEE J. Robotics Autom.*, vol. 3, pp. 361–373, Aug. 1987.
[16] Y.C. Kim and J.K. Aggarwal, "Determining object motion in a sequence of stereo images," *IEEE J. Robotics Autom.*, vol. 3, pp. 599–614, Dec. 1987.
[17] N. Ayache and B. Faverjon, "Efficient registration of stereo images by matching graph descriptions of edge segments," *Int'l J. Comp. Vision*, pp. 107–131, 1987.
[18] U.R. Dhond, "Stereo image interpretation in the presence of narrow occluding objects." PhD thesis, University of Texas at Austin, Department of Electrical and Computer Engineering, Austin, Tex., USA, Aug. 1992.
[19] J.J. Little and W.E. Gillett, "Direct evidence for occlusion in stereo and motion," *Img. Vision Computg.*, vol. 8, pp. 328–340, Nov. 1990.

Direct Extraction of Topographic Features for Gray Scale Character Recognition

Seong-Whan Lee and Young Joon Kim

***Abstract*—Optical character recognition(OCR) traditionally applies to binary-valued imagery although text is always scanned and stored in gray scale. However, binarization of multivalued image may remove important topological information from characters and introduce noise to character background. In order to avoid this problem, it is indispensable to develop a method which can minimize the information loss due to binarization by extracting features directly from gray scale character images.**

In this paper, we propose a new method for the direct extraction of topographic features from gray scale character images. By comparing the proposed method with Wang and Pavlidis' method, we realized that the proposed method enhanced the performance of topographic feature extraction by computing the directions of principal curvature efficiently and prevented the extraction of unnecessary features. We also show that the proposed method is very effective for gray scale skeletonization compared to Levi and Montanari's method.

***Index Terms*—Gray scale character recognition, principal curvature, principal orthogonal elements, topographic feature extraction.**

I. Introduction

It is well known that the character images of documents seen by the scanners have a wide range of gray scale for the following reasons [1]:

- Gray or different color background of documents as in magazine and business forms
- Textured background of documents as in bank checks
- Different types of ink
- Convolution distortion because of the point spread function of the scanner
- Nonuniform illumination of the scanner
- Multiplicative noise such as nonuniform paper reflection
- Additive noise due to electronics

When such a gray scale character image is binarized there is a significant information loss: Typically areas which are narrow compared to the support of the point spread function disappear, resulting in broken characters or touching characters.

Two kinds of methods have been used to overcome the disadvantages of simple binarization such as thresholding [1]. The first is to do binarization in a far more careful way than it is currently done in OCR. A representative of this kind of method is adaptive thresholding [2]. However, it requires a relatively large amount of computation, and no matter how good the binarization is, it cannot completely avoid the loss of information. The second approach is to perform recognition without binarization by using techniques related to matched filters [3], [4]. However, this approach is limited to applications such as recognition of single font characters where there is only a limited variation on the form of the characters.

A more promising approach is to perform feature extraction directly from gray scale character images [1]. By doing so, it bypasses the phase of binarization and thus avoids the loss of information caused by the operation.

In this paper, we propose a new method for extracting topographic

Manuscript received Mar. 8, 1994; revised Dec. 26, 1994.

S.-W. Lee is with the Dept. of Computer Science, Korea University, Seoul 136-701, Korea; e-mail: swlee@ailab.korea.ac.kr.
Y.J. Kim is with BIT Computer Co., Ltd., 818-11, Yeoksam-dong, Kangnam-ku, Seoul 135-080, Korea.

IEEECS Log Number P95072.